Title: Understanding the optical responses of leaf nitrogen in Mediterranean Holm oak (*Quercus ilex*) using field spectroscopy.

Authors: Javier Pacheco-Labrador [1], Rosario González-Cascón [2], M. Pilar Martín [1], David Riaño[1, 3]

[1] Laboratorio de Espectro-radiometría y Teledetección Ambiental (SpecLab), Instituto de Economía, Geografía y Demografía, Centro de Ciencias Humanas y Sociales, Consejo Superior de Investigaciones Científicas (CSIC), C/Albasanz 26-28, 28037Madrid,Spain

[2] Departamento de Medio Ambiente, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Ctra. Coruña, Km. 7,5, 28040 Madrid, Spain

[3] Center for Spatial Technologies and Remote Sensing (CSTARS), Department of Land, Air and Water Resources, University of California, One Shields Avenue, Davis, CA 95616-8617, USA

e-mails: javier.pacheco@cchs.csic.es; cascon@inia.es; mpilar.martin@cchs.csic.es; david.riano@cchs.csic.es

Current address / Correspondence address:
Javier Pacheco-Labrador
Laboratorio de Espectro-radiometría y Teledetección Ambiental (SpecLab),
Instituto de Economía, Geografía y Demografía
Centro de Ciencias Humanas y Sociales
Consejo Superior de Investigaciones Científicas
C/Albasanz 26-28
Madrid, 28037
Spain
Phone: +0034 91-602-27-84
Fax: +0034 91-602-29-71

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

ABSTRACT

The direct estimation of nitrogen (N) in fresh vegetation is challenging due to its weak influence on leaf reflectance and the overlaps with absorption features of other compounds. Different empirical models relate in this work Leaf Nitrogen Concentration ($[N]_{Leaf}$) on Holm oak to leaf reflectance as well as derived spectral indices such as Normalized Difference Indices (NDIs), the Three Bands Indices (TBIs) and indices previously used to predict leaf N and chlorophyll. The models were calibrated and assessed their accuracy, robustness and the strength of relationship when other biochemicals were considered. Red Edge was the spectral region most strongly correlated with $[N]_{Leaf}$, whereas most of the published spectral indexes did not provide accurate estimations. NDIs and TBIs based models could achieve robust and acceptable accuracies ($TBI_{1310,1720,730}$: R^2 = 0.76, [0.64,0.86]; RMSE (%) = 9.36, [7.04,12.83]). These models sometimes included indices with bands close to absorption features of N bonds or nitrogenous compounds, but also of other biochemicals. Models were independently and inter-annually validated using the bootstrap method, which allowed discarding those models non-robust across different years. Partial correlation analysis revealed that spectral estimators did not strongly respond to $[N]_{Leaf}$ but to other leaf variables such as chlorophyll and water, even if bands close to absorption features of N bonds or compounds were present in the models.

Keywords: Field Spectroscopy, Leaf Nitrogen, *Quercus ilex*, Spectral index

1. INTRODUCTION

Understanding how matter is exchanged through atmosphere, vegetation and soil is fundamental in assessing how far human activities can alter biochemical natural

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

cycles and impact climate change. Natural equilibriums in the nitrogen cycle have been dramatically broken as a consequence of human activities, with important impacts on terrestrial and aquatic ecosystems (Galloway, 1998; Keeney and Hatfield, 2001; Sutton, et al., 2011). The nitrogen cycle is also related to Global Warming, though has two opposing effects: on one hand, NO_2 is a greenhouse gas which directly and indirectly contributes to radiative forcing; on the other hand, nitrogen (N) has allowed the CO_2 sink increase, thus existing chances for a "smart management" linking both N and carbon (C) cycles (Sutton, et al., 2011). Increases of atmospheric carbon concentration and N available for vegetation have triggered changes in Gross Primary Productivity (GPP), since both compounds are inputs required in the photosynthetic process, affecting also species distribution and ecosystems composition (Beier, et al., 2008; Cheng, et al., 2009; Langley, et al., 2009; Sardans, et al.; Sutton, et al., 2011).

Different approaches have been used to estimate vegetation N contents by means of laboratory and field spectroscopy, and airborne and satellite hyperspectral remote sensing data. Some authors have tried to overcome local variability in heterogeneous environments such as Mediterranean ecosystems (Serrano, et al., 2002); others have looked for generalizable spectral indices by studying species from different ecosystems (Ferwerda, et al., 2005; Martin, et al., 2008); and others have used multi-temporal data in order to increase available N ranges and relate them with optical response changes (Tian, et al., 2011; Yao, et al., 2010). Though remote sensing has shown to be able to quantify N at different scales, the most relevant spectral regions reported in the literature have usually been related to chlorophyll absorptions and mesophyll scattering in the Near-Infrared (NIR) (Ferwerda, et al., 2005; Martin, et al., 2008; Tian, et al., 2011; Yao, et al., 2010). Estimations based on wavelengths located near N bonds and protein absorption bands have also been reported, showing sometimes high coefficients

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

of determination (Martin, et al., 2008; Ourcival, et al., 1999; Serrano, et al., 2002; Smith, et al., 2003). However, Jacquemoud, et al. (1996) questioned that relationships between N or proteins and spectral reflectance in the NIR and Short Wave Infrared (SWIR) were actually caused by the optical response of these molecules and N bonds, being likely that relations found were due to covariance with other more abundant compounds such as water or pigments.

Though usually up to 6% of dry leaf mass is attributable to N (Bergmann, 1992; Jacquemoud, et al., 1996; Wright, et al., 2004), water could reduce the influence of N optical responses in fresh vegetation (Fourty, et al., 1996; Jacquemoud, et al., 1996; Kokaly, 2001; Schlerf, et al., 2010). Even though absorptions related to N bonds (and others) are found in the NIR and SWIR (Burns and Ciurczak, 2008; Curran, 1989; Fourty, et al., 1996), they are harmonics and overtones of fundamental absorptions and combination bands (Burns and Ciurczak, 2008; Curran, 1989; Fourty, et al., 1996). These bands are broad (half-depth bandwidth 30-40 nm) and much weaker than their corresponding fundamentals (Burns and Ciurczak, 2008; Kokaly, 2001). Besides, in the optical domain appear several absorption bands: harmonics and overtones corresponding to H-C, H-O, C-O bonds and also other important functionalities. This makes that location, width and height of all these bands vary due to adjacent and overlapping bands, not being possible to isolate and directly relate any of these bands with a single biochemical (Burns and Ciurczak, 2008; Curran, 1989; Osborne and Fearn, 1986). Furthermore, N is present in diverse leaf components which show different optical responses not always related with the N bonds. N can be partially found in leaf pigments, especially in chlorophyll-a (6.27% N mass) and chlorophyll-b (6.18% N mass) (Nelson, et al., 2004), whose strong optical responses in the Visible spectral region are due to single and double carbon–carbon bonds found in phytol tails (Nelson,

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

et al., 2004). N atoms are located in the porphyrin ring, surrounding a central magnesium ion and not forming amide bonds (usual in proteins), but C-N bonds, which show an absorption feature around 2180 nm (Burns and Ciurczak, 2008; Curran, 1989; Fourty, et al., 1996). N is also present in proteins (typically16% N mass), especially in the ribulose-1,5-biphosphate carboxylase-oxygenase (Rubisco), where percentages between 30–50% of the N in green leaves have been found (Elvidge, 1990). Proteins show absorption features not only related with N bonds, but with C, H and O in the SWIR, sometimes also present in other molecules such as cellulose, lignin, etc. (Burns and Ciurczak, 2008; Curran, 1989; Fourty, et al., 1996). Finally, contents and proportions of nitrogenous compounds in the leaf can change along time in response to phenology and environmental conditions (Sabaté, et al., 1995; Silla, et al., 2010), doing so their contributions to leaf reflectance.

Leaf N has usually been estimated using spectral bands related to chlorophyll absorption features and another compounds and leaf characteristics such as the NIR scatter in the spongy mesophyll (Hansen and Schjoerring, 2003; Reddy and Matcha, 2010; Tian, et al., 2011). However, chlorophylls actually represent a reduced portion of the total leaf N (Evans, 1983; Evans, 1989; Yoder and Pettigrew-Crosby, 1995), and the relationships found are species-dependent (Ferwerda, et al., 2005). Different analyses have been used to select spectral information able to predict leaf N, some examples are shown in Table 1. Though some of the mentioned works achieved high coefficients of determination, the involved spectral bands were different in each case, and not always close to absorptions related to N bonds.

INSERT TABLE 1 HERE

This work investigates the influence of the optical responses of foliar N on the estimation of leaf nitrogen concentration ($[N]_{Leaf}$) in Holm oak (*Quercus ilex* L. *subsp.*

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

rotundifolia) using field spectroscopy. Since a), leaf optical response is driven by its biochemical constituents and structure, being the signals related to N bonds weak in the optical domain (Burns and Ciurczak, 2008; Jacquemoud, et al., 1996); and b) plant constituents content and structure vary over the vegetative year driven by environmental conditions and the leaf maturation process (Correia, et al., 1992; Crescente, et al., 2002; Milla, et al., 2010; Pinto, et al., 2011); we test the robustness of the empirical models analyzing the influence of different environmental conditions on the leaf traits and reflectance by using data from different vegetative years. Then we look for estimators based on absorption features of N bonds and nitrogenous compounds such as proteins or chlorophyll, and analyze the strength of $[N]_{Leaf}$ in the predictive models when other constituents of the leaf are considered.

3. METHODOLOGY

3.1 Study Site

The study site is located in Las Majadas del Tietar, Cáceres (Spain) (39°56‘29" N, 5°46'24'' W) within a typical Mediterranean ecosystem called “dehesa”,with a long history of human management, which covers large areas in Western Spain and Portugal (2,200,000 ha, ~3.77% of the Iberian peninsula surface) (Blanco, et al., 2005;Pulido García, 2002). A flux tower is operated at the site by the Centro de Estudios Ambientales del Mediterráneo (CEAM) since May 2003 to measure water and carbon fluxes using the eddy-covariance method and included in the FLUXNET (fluxnet.ornl.gov) and SpecNet (http://www.specnet.info) international networks (Baldocchi, 2003).

Vegetation structure is a wooded grassland or savanna, where grazing pasture is mixed with scattered trees, mostly Holm oak. This is a typical Mediterranean species

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

which can be found in savanna ecosystems as well as in denser forests and mixed communities, mainly in the Mediterranean basin. In our study area, trees cover about 20% of the ground and are on average 14 m apart from each other. Mean annual rainfall is 572 mm and mean annual temperature is 16.7 ºC, ranging between monthly averages of 1.4 ºC and 34.9 ºC. The summer is hot, with maximum temperatures beyond 40 °C, and dry. Most of the rain occurs during the winter and the beginning of the spring that coincide with the spring of new Holm oak leaves. At this early leaf stage, chlorophyll content is low but increases throughout its first year, though summer drought temporally pauses this trend (Silla, et al., 2010). Leaves start to become leathery after summer developing some defensive structures as they turn into completely mature leaves to stay for at least one more year on the branches (Morales, et al., 2002).

3.2 Nitrogen concentration, other biophysical variables and spectral data

Holm oak leaves from current season shoots were collected in a monthly basis from April 2009 to March 2011. Ten random trees were sampled during the first sampling campaign within an area of 35.3 ha that, after April 2010, were reduced to five trees, as they were still representative of the spatial variability within the study area. Each tree crown was divided in quarters; upper-lower and north-south sections and leaves were sampled from each quarter. For N content determination, at least 100 new leaves per tree that sprouted that year were sampled and dried 48 hours at 65 ºC in the laboratory. Later, unwashed, grinded foliar samples were analyzed with the dry combustion method to determine carbon (C) and N using a LECO instrument (CN 2000) (Montoya Moreno, et al., 1997).

Moreover, 12 leaves per tree quarter were sampled. From which three leaves per quarter were selected and measured with a SPAD-502 Leaf Chlorophyll Meter

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

(http://www.konicaminolta.com). Over the same leaves, leaf reflectance was measured with an ASD FieldSpecTM 3 spectroradiometer (www.asdi.com). The fiber optic was connected to a Plant Probe with a Leaf Clip supplied by the same company. Reflectance spectra were limited from 400 to 2500 nm. To avoid biases related to illumination changes and dark current drift, spectroradiometer and probe bulb were turned on for 30 minutes before measuring; moreover optimization, calibration and dark current correction were frequently performed during the measurements. Quality control was applied to the data in order to eliminate those spectra acquired when leaves did not completely cover the Field of View (FOV) of the spectroradiometer.

The 12 leaves selected per quarter of each tree (including those measured with SPAD and the spectroradiometer) were cooled until analysis in the laboratory to calculate water (C_w, g/cm^2) and dry matter (C_m, g/cm^2). Fresh and dry weights were measured by gravimetric method whereas leaf area was estimated by leaf scanning using an Epson Perfection V30 color scanner (www.epson.com) (Viegas, et al., 1992). A k-means clustering algorithm implemented in PCI image analysis software (www.pcigeomatics.com) was used to calculate leaf area automatically. Multiplying leaf N and C contents in mg/g, by C_m, $[N]_{Leaf}$ and leaf carbon concentration, ($[C]_{Leaf}$) in mg/cm^2 were finally calculated.

Chlorophyll concentration (C_{a+b}, μg/cm^2) was estimated from the SPAD values using a model calibrated in the laboratory from additional samples. These were acquired in four dates along the campaign 2010-2011 in the same five trees used for routine measurements. Six leaves per tree from different orientations in the canopy were sampled. After field SPAD measurements, each leave was individually packed in black plastic bags and stored cold in darkness. Samples were kept frozen (-20 ºC) until analysis. Chlorophyll analyses were performed in the laboratory with the DMSO

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

extraction method (Silla, et al., 2010; Wellburn, 1994) and SPSS 17 statistics software (www.ibm.com/SPSS_Statistics) was used to adjust the model. Eventually leaf spectral reflectance, C_{a+b}, C_w, C_m, $[N]_{Leaf}$ and $[C]_{Leaf}$ were averaged by date and tree.

3.3 Estimating leaf nitrogen concentration with field spectroscopy

We calibrated empirical models to estimate $[N]_{Leaf}$ from field spectra using data acquired from 2009 young leaves (68 data), whereas samples belonging to the 2010 dataset were used for model validation (50 data). Linear regression was established between $[N]_{Leaf}$ and leaf reflectance (interpolated to 1nm step by the radiometer) from 400 to 2500nm. We also calibrated linear models using a selection of 82 indices proposed/used in the literature for chlorophyll and/or N estimation. Finally, Normalized Difference Indices (NDIs) and Three Bands Indices (TBIs) (Eq. 1 and 2 respectively) were generated using all the possible spectral band combinations within the 400-2500 nm range. In both cases, a pre-selection of bands was necessary in order to reduce the computational effort. One out of five bands was selected to generate a total of 177,241 NDIs whereas one out of ten was used to calculate 9,393,931 TBIs. All the calculations were implemented in Matlab 7.9 (http://www.mathworks.com).

$$NDI_{i,j} = \frac{\rho_i - \rho_j}{\rho_i + \rho_j};\ \forall i,j \in [400,2500] \quad (1)$$

$$TBI_{i,j} = \frac{\rho_i}{\rho_j + \rho_k};\ \forall i,j,k \in [400,2500] \quad (2)$$

In order to test the models robustness, the Bootstrap non-parametric analysis was applied (Efron, 1979). This technique allows calculating confidence intervals for any estimated statistic avoiding any assumption about the probability distribution of the

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

data. Models were calibrated using random subsamples of the calibration dataset selected with replacement. The 95% confidence intervals of every statistic (coefficients of determination, p-values, errors, etc.) were established from the percentiles. Thereby, it was possible to know whether the statistics estimated in the validation process were different from those predicted by the model built using the calibration dataset. Bootstrap technique has proved to perform better than other analyses used for error or sensibility estimation of empirical models, such as jackknife or cross-validation (Efron, 1979). For every model, 20 subsamples of 50 data were generated.

Independent validation of models calibrated with the 2009 data was accomplished using the 2010 dataset. This way we tested the robustness of the models using different leaves (independent validation) sampled under the same environmental conditions, but also we assessed the inter-annual performance of the models which implies considering potential variation in these environmental conditions. Models were rejected when Root Mean Squared Error (RMSE) (%) for the validation dataset were out of the confidence intervals of the calibration RMSE (%). On the contrary, those models that passed this test were considered inter-annually consistent and suitable for $[N]_{Leaf}$ estimation.

3.4 Models explanation. Partial correlation analyses

We used partial correlation analysis to measure the intensity of the linear relationships between $[N]_{Leaf}$ and the spectral variables in the empirical models while taking into account the inlfuence of other compounds in the leaf. For that, we selected biophysical variables known to have significant influence on the leaf spectral response (Feret, et al., 2008; Jacquemoud, et al., 1996): C_w (g/cm^2), C_m (g/cm^2), and C_{a+b} ($\mu g/cm^2$). We also used $[C]_{Leaf}$ as a proxy of the foliar structural (cellulose, lignin) and

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

nonstructural (starch) polysaccharides concentrations, since several of their absorption features are close to those related to N bonds and proteins (Burns and Ciurczak, 2008; Curran, 1989; Fourty, et al., 1996).

We generated a model where the spectral variables, the predictors of $[N]_{Leaf}$, were a linear combination of the selected biophysical variables (C_w, C_m, C_{a+b} and $[C]_{Leaf}$) and $[N]_{Leaf}$. Then, two different statistics were computed: partial correlation coefficients and fractions (details can be found in Borcard, 2002; Legendre and Legendre, 1998a; Legendre and Legendre, 1998b). Partial correlation coefficients measure the proportion of the variance that is explained by the addition of one biophysical variable to the model while the rest of the variables are held constant with respect to the dependent (spectral) and the analyzed ones. It explains the exclusive mutual relationship between both. Fractions measure the proportion of the variance explained by one biophysical variable when the other variables are held constant with respect to this one but not to the dependent. It is the part of the coefficient of multiple determination (R^2) explained only by the variable analyzed (called fraction [a]). Moreover, we calculated the proportion of the variance explained by the rest of the variables (fraction [c]), by the combination of all the variables (the coefficient of multiple determination R^2) regarding the substracted fractions [a] and [c], (the fraction [b]), and the unexplained proportion [d].

Partial correlation coefficients were calculated for all the biophysical variables, whereas fractions were computed only for $[N]_{Leaf}$. We applied partial correlation analysis to leaf reflectance, the 82 spectral indices found in the literature, and the validated NDIs and TBIs.

4. RESULTS

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

4.1 Temporal behavior of leaf nitrogen concentration and optical response

Holm oak $[N]_{Leaf}$ leaves showed a similar behavior along the two vegetative years. $[N]_{Leaf}$ increased from approximately 0.25 mg/cm^2 in spring to about 0.35 mg/cm^2 a year later (Fig. 1). After that, these leaves (already mature) slightly reduced their $[N]_{Leaf}$ at the end of the vegetative period, remaining then stable (not shown). Average $[N]_{Leaf}$ was in the range registered for Holm oak under Mediterranean climatic conditions (Gratani, et al., 2006; Gulías, et al., 2002; Limousin, et al., 2010; Peña-Rojas, et al., 2005). C_{a+b} evolved similarly but the increase slowed down during the summer; C_m, C_w and $[C]_{Leaf}$ initially increased and remained stable after the summer, during which C_w slightly decreased.

INSERT FIG 1 HERE

Holm oak leaves exhibited a differential behavior depending on the spectral region. As can be seen in Fig. 2, reflectance in Visible bands was initially high in the spring, decreasing slightly until it started peaking again in the summer, to largely decrease in the fall. Red Edge and NIR reflectance evolved similarly, but NIR also increased in the early spring. The SWIR region experienced little changes during the whole period, except a noticeable increase in August 2009.

The two vegetative years studied showed different environmental conditions. The first (2009-2010) was drier and colder, with spring and autumn precipitations of 72 and 274 mm respectively and mean annual temperature of 15.6ºC. During the second one (2010-2011) precipitations were doubled (196 and 418 mm) and mean annual temperature slightly higher 17.0ºC.

INSERT FIG 2 HERE

4.2 Leaf nitrogen concentration and spectral reflectance bands

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

Fig. 3 shows coefficients of determination and their 95% confidence intervals between leaf reflectance and $[N]_{Leaf}$. The Green, NIR and Red Edge spectral regions achieved the highest R^2, which were low in the SWIR and minimum around the water absorption bands at 1390 nm, 1900 nm and 2500 nm. Maximum R^2 values were found at 729 nm (R^2 = 0.69) and 540 nm (R^2 = 0.64), with broad confidence intervals. The statistic p-value also showed wide intervals, the upper limit was larger than 0.05 from 1074 to 2500nm, where the relationships between $[N]_{Leaf}$ and reflectance were not significant. RMSE ranged between 10.77% and 19.28%, with confidence intervals around -3.48% and +2.31% in average. Errors in $[N]_{Leaf}$ estimated with the validation dataset were mostly out of the confidence interval, except in the range 724 – 1172 nm.

INSERT FIG 3 HERE

4.3 Leaf nitrogen concentration and published spectral indices

Table 2 presents the statistics of the empirical models built to relate the spectral indices designed for chlorophyll and/or N estimation selected from the literature and Holm oak $[N]_{Leaf}$. The highest coefficients of determination were achieved by the indices B(log(1/ρ730))' (R^2 = 0.64) proposed by Yoder and Pettigrew-Crosby (1995) for chlorophyll and N estimation; secondly ρ860/(ρ550 · ρ708) (R^2 = 0.59) and ρ672/(ρ550 · ρ708) (R^2 = 0.57), both proposed by Datt (1998) for pigments estimation; and also $NDI_{1220,710}$ (R^2 = 0.57), proposed by Zhu, et al. (2007), for N estimation in rice. Only 16 indices achieved correlations with average R^2 larger than 0.50, though no one left this value out of the 95% confidence interval. 71 indices showed average p-values lower than 0.05, but these were reduced to 42 when 95% confidence intervals were considered. RMSE (%) values were inversely proportional to the coefficients of determination, ranging between 11.53% and 19.29% in average. However, errors

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

estimated with the validation dataset (2010) were most of the times out of the 95% confidence interval of calibration RMSE (%), showing lower values. As a consequence, only 11 models could be validated and only 7 of them were significant.

INSERT TABLE 2 HERE

4.4 Leaf nitrogen concentration and Normalized Difference Indices

INSERT FIG 4 HERE

Fig. 4 shows the R^2 values for all spectral bands combined to build the NDIs. The strongest correlations were reached combining reflectance at NIR bands with Green and Red Edge or other NIR bands. R^2 was low for those indexes that included SWIR bands; however values locally larger were achieved when close SWIR bands and/or water absorption bands were combined. Some of these were close (±10nm distance) to absorption features of N or other atoms bonds (Burns and Ciurczak, 2008; Curran, 1989; Fourty, et al., 1996), or to in-vivo measured absorption peaks of proteins, cellulose + lignin or water (Jacquemoud, et al., 1996).

INSERT FIG 5 HERE

The 2000 NDIs showing the largest R^2 were selected, confidence intervals of the models statistics were estimated by bootstrap and the models were validated with the independent dataset from year 2010 (Fig. 5). The band i of the index (Eq. 1) was found around six spectral regions 540, 730, 825, 905, 1050 and 1200 nm, where the coefficients of determination were locally maximum. Band j (Eq. 1) was broadly distributed between 780 and 1380 nm. Independent validation allowed selecting 722 of the 2000 indices. As can be seen in Fig. 5, all the indices including Green bands and most of the indices where Red Edge bands participated could not be validated. On the contrary, only indices where band i was above 735 nm passed the validation. In the case

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

of band j, validation reduced the dispersion of locations found in the NIR, remaining narrower bands.

INSERT FIG 6 HERE

Validated NDIs were classified by the different bands that compose them in Fig. 6, with their respective RMSE (%). Indices whose band i was below 1000 nm, such as those based on Red-Edge, reached lower R^2 than those with band i around 1050 or 1100 nm; and unlike these, showed validation errors not centered, but close to the low bound of the 95% confidence interval. The indices with the highest R^2 of each band combination are summarized in Table 3.

INSERT TABLE 3 HERE

4.5 Leaf nitrogen concentration and Three Bands Indices

INSERT FIG 7 HERE

From the 2000 TBIs with the highest R^2, 1479 could be validated (Fig. 7). The validation process discarded indices including combinations of all bands, and always reaching $R^2 < 0.755$. The band i, the numerator of the indices (Eq. 2), was composed by bands centered around 1160 and 1310 nm. The denominator resulted from the addition of reflectances from a narrow band in the Red Edge, around 730 nm (band k), and a wide range of bands all along the SWIR region (band j) except in the water absorption regions. Band k showed six local maxima of R^2 at 1570, 1630, 1720, 1810, 2180 and 2300 nm.

INSERT FIG 7 HERE

The combinations of bands found in the 1479 validated TBIs and their statistics are summarized in Fig. 8. Indices whose band i was located around 1160nm casted the lowest R^2 (Table 4). The highest determination coefficients were found in those indices

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

with band i around 1310 nm and band j within SWIR-1 (1500-1900 nm). Validation errors were always found close to the low bond of the 95% confidence interval.

INSERT TABLE 4 HERE

4.6 Models explanation. Partial correlation analyses

For leaf reflectance, $[N]_{Leaf}$ partial correlation coefficients were larger in the Visible and NIR regions than for other variables, except C_{a+b}, highly related to Green and Red Edge bands (Fig. 9). SWIR reflectance is mainly related to C_w below 1870nm, whereas $[N]_{Leaf}$ was better correlated than the rest of variables above this point. $[C]_{Leaf}$ partial correlation coefficients were low in the whole spectrum and for all the spectral indices analyzed. However, fractions showed that a great part of the variance of reflectance bands remained unexplained (fraction [d]), excluding Green and Red Edge regions, where variance was explained by [b], the combination of all the variables together removed fractions [a] and [c].

In the case of spectral indices selected from literature, the highest partial correlation coefficients (around 0.2) were for $[N]_{Leaf}$, C_{a+b} and C_w. However, when fractions were analyzed, most of the variance was explained by [b] or not explained ([d]). Partial correlation coefficients of NDIs and TBIs were higher for chlorophyll than for the rest of variables in all cases. The second largest coefficients were for $[N]_{Leaf}$ and C_w, in the case of NDIs, and only C_w in the case of TBIs. For both indices, the fraction [b] was about 60-70% and unexplained variance ([d]) was slightly larger than [c] the most of the times. The fraction of variance exclusively explained by $[N]_{Leaf}$ ([a]) was minimum, and lower than 10%.

INSERT FIG 9 HERE

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

5. DISCUSSION

5.1 Temporal behavior of leaf nitrogen concentration and optical response

The biophysical variables concentrations increased along the first year of development of the Holm Oak leaves. Summer drought slowed some of these increases, which recovered after the season. C_{a+b} increase was perceived as a change in leaf color from light to dark green. Moreover, after summer, leaves became leathery, developing waxes and other defensive structures.

Holm oak leaves optical responses were coherent with the observed changes in biophysical variables. On one hand, physiological characteristics of this sclerophyll species, well adapted to drought, produced low variations in the SWIR region (an also in C_w). The most noticeable change in the study period was one episode of extreme drought which occurred in August 2009. In the Visible region, reflectance seemed controlled by the increase of chlorophyll concentration. NIR reflectance increased along spring and summer with the development of the young leaves, and decreased in autumn, when leaves acquired a more leathery structure, more similar to mature ones. The behavior of biophysical variables and optical responses were slightly different in each vegetative year, as environmental conditions were. This inter-annual variability may modify the relationships between biophysical and spectral variables every year.

5.2 Leaf nitrogen concentration and spectral reflectance bands

The highest coefficients of determination were found in the Green, the Red Edge and the NIR regions. This may be partially explained by the changes in C_{a+b} and the development of leaf structure, which greatly covaried with $[N]_{Leaf}$. The Red Edge experimented the largest variation along the year and achieved also the largest coefficients of determination with $[N]_{Leaf}$. The Red Edge results from the contrast

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

between Red absorption of chlorophylls and the intense NIR scatter into the spongy structure of the leaves mesophyll (Seager, et al., 2005).

Water absorption bands were completely uncorrelated to $[N]_{Leaf}$. The slight changes showed by SWIR leaf reflectance and the decouple of precipitation regimen and $[N]_{Leaf}$ may explain this fact. In the SWIR region minor local maxima of R^2 were found, some of them close to the protein absorption peaks around 1730 nm (C-H Stretch), 2180 nm (N-H bend 2nd overtone, C-H stretch/C=O stretch or C=O stretch/ Amide III^{b} combinations), 2300 nm (C-H bend second overtone) and 2352 nm (C-H_2 bend second overtone) . However, these maxima were also close to absorptions at 1730 nm (C-H stretch first overtone), 2100 nm (O-H bend/C-O stretch combination, Asym. C-O-O stretch third overtone), and 2335 nm (C-H stretch / C-H deformation), related to lignin and cellulose (Burns and Ciurczak, 2008; Curran, et al., 1992; Fourty, et al., 1996; Jacquemoud, et al., 1996). Another observed local maximum cannot be related to any absorption feature of N or proteins. Instead, it is impacted by a lignin + cellulose feature at 1200 nm (O-H bend first overtone) (Burns and Ciurczak, 2008; Curran, 1989; Jacquemoud, et al., 1996), This may be explained by covariation between lignin and N (Schlerf, et al., 2010). In all the cases, there were shifts between the maxima of correlation and the absorption peaks, always away from the water bands which showed no correlation with $[N]_{Leaf}$. Thus water absorptions may be masking another known absorption features.

Validation discarded all the models except those within the range 724-1172nm. These models did not always show the highest coefficients of determination and lowest calibration errors. Other bands strongly related to $[N]_{Leaf}$, such as the Green were discarded.

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

5.3 Leaf nitrogen concentration and published spectral indices

Chlorophyll and $[N]_{Leaf}$ spectral indices found in the literature generally failed to predict $[N]_{Leaf}$ in Holm oak, showing non-significant relations or being discarded in the validation process. The index with the highest coefficients of determination, the B(log(1/ρ730))', is based on a single Red Edge spectral band. The Red Edge band and the index model statistics are equal, thus we can conclude in this case that the transformation of reflectance did not improve the $[N]_{Leaf}$ estimation. Some indices such as the one proposed by Ferwerda, et al. (2005), empirically developed to estimate N in different species showed low correlations (R^2 = 0.21). As the author recognizes, this index is governed by chlorophyll and cellulose absorption features and not by N bond or protein absorptions. Other indices developed under a more physical approach such as the NDNI Serrano, et al. (2002), or the one proposed by Zhao, et al. (2006) based on the 1510 nm and 2054 nm N absorption features respectively were completely uncorrelated with our $[N]_{Leaf}$ dataset (R^2 = 0.01 both). Additionally other indices proposed by Tian, et al. (2011) based on Red Edge and Blue bands (R^2 > 0.80 in rice) were neither correlated to Holm oak $[N]_{Leaf}$ (R^2 < 0.22). On the other hand the MERIS Terrestrial Chlorophyll Index (MTCI) achieved a significant and robust (validated) relationship, though R^2 was low (R^2=0.42). This index is operationally used for chlorophyll content estimation and will be also used in the Sentinel missions (Dash and Curran, 2004). Regarding the spectral regions included in the indices no clear patterns were found, though Red Edge and Green bands were usually involved in the highest correlated indices.

5.4 Leaf nitrogen concentration and Normalized Difference Indices

Modest improvement in $[N]_{Leaf}$ estimation is achieved by NDIs when compared with the estimations obtained using single band at 730 nm. Mean coefficients of

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

determination and errors of the best predictors are statistically equal. NDIs based on Green and Red Edge bands did not reach higher coefficients of determination than the single spectral bands that were used to compute them. Actually, most of the indices based on these bands were discarded in the validation process.

NDIs included bands close to absorptions related to oils, water, cellulose, starch and lignin (Table 3) (Burns and Ciurczak, 2008; Curran, 1989; Fourty, et al., 1996). Only $NDI_{915,920}$, showed bands close of an absorption feature related to proteins (a C-H stretch third overtone) (Curran, 1989). The selection of NDIs with bands at 1120 and 1200nm may be explained by covariation between lignin and N (Schlerf, et al., 2010). On the other hand, no absorption peaks have been measured in-vivo in these spectral regions (Jacquemoud, et al., 1996). Local maxima of correlation found in the SWIR region may suggest detection of weak optical responses of N bonds. However bands close to N bonds or protein absorptions were also near to cellulose or lignin absorption features. Only in the case of the $NDI_{1395,2180}$ the SWIR band had been related to N bonds or proteins exclusively (Burns and Ciurczak, 2008; Curran, 1989).

Other authors that have also tested NDIs for N estimation (Tian, et al., 2011; Yao, et al., 2010) found high coefficients of determination when bands in the Red Edge and NIR or Blue regions were combined in the indices, being thus likely related to chlorophyll and leaf structure. Yao, et al. (2010), using canopy reflectance measurements on wheat and combining spectral data into NDIs with bands at 720 nm and between 930 and 1010 nm, achieved high correlations ($R^2 > 0.80$). This combination is similar to one of the indexes we validated, the $NDI_{730,960}$. Red Edge reflectance does not show any direct relationship with N absorptions, though is mainly result of chlorophyll absorptions and mesophyll scatter in the NIR (Seager, et al., 2005). On the other hand, three absorptions related to starch, oil and water appear in the 930–

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

1010nm interval (Burns and Ciurczak, 2008; Curran, 1989). Thus NDIs selected in this work may not actually respond to N bond or protein absorptions, but maybe chlorophylls when Red Edge is involved; and also with lignin, other organic compounds or the structural development of the leaf.

5.5 Leaf nitrogen concentration and Three Bands Indices

Bands with optical responses dominated by different compounds were mixed in the selected TBIs. The numerator of the indices (band i) was centered around 1160 and 1310 nm, these bands showed low correlation with $[N]_{Leaf}$, being the first in a local minimum of correlation and the second completely uncorrelated (Fig.3). None of the bands are related to protein or cellulose + lignin absorption features (Jacquemoud, et al., 1996). The 1160nm band responds to C=O stretch, fourth overtone absorption, but not related to any particular compound (Burns and Ciurczak, 2008). The Red Edge was always included in the denominator (band k), not being directly related neither to protein nor N bond absorption features. This region is governed instead by chlorophyll absorption and NIR scatter (Seager, et al., 2005). The second band in the denominator (band j) included bands all along the SWIR region, but they were water absorption bands instead of N. The six local maxima of correlation found in SWIR were close to reported N absorptions of several bonds (Burns and Ciurczak, 2008; Curran, 1989; Fourty, et al., 1996) and absorption features of proteins and / or cellulose + lignin (Jacquemoud, et al., 1996) (Table 4). Again, only the one at 2180 nm could be related to N and protein absorption features exclusively (Burns and Ciurczak, 2008; Curran, 1989; Kokaly, 2001). Selected TBIs gathered an uncorrelated band as a reference, and combined variations of Red Edge reflectance (R^2 = 0.69) with reflectance at SWIR wavelengths with a weak relationship but locally maximum in some cases. TBIs

Published in International Journal of Applied Earth Observation and Geoinformation, Volume 26, February 2014, Pages 105-
http://dx.doi.org/10.1016/j.jag.2013.05.013

selected in previous works (Tian, et al., 2011) were based on Red Edge, NIR and Blue bands, whereas those here selected were based on Red Edge in combination with NIR and SWIR and Red Edge bands.

Validated and selected TBIs reached the highest average coefficients of determination and the lowest average errors of all the $[N]_{Leaf}$ estimators here tested ($TBI_{1310,1720,730}$: R^2=0.76; [0.64,0.86]; RMSE (%)=9.36; [7.04,12.83]). Unlike in the case of NDIs, the most correlated models were also the most robust. However, looking at the confidence intervals, there are no differences between TBIs and NDIs statistics (Tables 3 and 4). This means that the models show the same R^2 and RMSE (%) values, though predictions would show different precision depending of the width of the intervals.

Each group of spectral variables was calculated selecting spectral bands with a different step. Original reflectance and derived published spectral indices were included in the models as provided by the spectroradiometer, interpolated at 1nm steps from the original data, even though spectral resolutions are coarser in VNIR and SWIR, (3 nm and 10 nm respectively) (ASD, 1999). On the contrary, NDIs and TBIs were built using one out of 5 or 10 bands in each case. The aim of this work was not to compare them, but the confidence interval widths in Fig. 3 and Tables 2, 3 and 4 suggest that no big differences would have been found if all the models had been built using the same spectral bands.

5.6 Models explanation. Partial correlation analyses

$[N]_{Leaf}$ showed, in general, low partial correlation coefficients and fractions. Fraction [b] showed negative values for reflectance bands between 1090 and 1515nm, 1800 and 2088nm and above 2381nm and also for some indices. These can be explained